DOCUMENT RESUME

ED 172 214 CS 204 866

AUTHOR Penfield, Elizabeth F.
TITLE Proficiency Testing within the Basic Writing Course: Developing a Faculty Consensus on Standards.
PUB DATE Apr 79
NOTE 11p.; Paper presented at the Annual Meeting of the Conference on College Composition and Communication (30th, Minneapolis, Minnesota, April 5-7, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Standards; College Freshmen; Compensatory Education; *Composition (Literary); Disadvantaged Youth; *Equivalency Tests; Higher Education; *Performance Based Education; *Program Development; *Remedial Programs; Student Ability; *Writing Skills

ABSTRACT

A proficiency test became the exit requirement for the noncredit, basic English course at the University of New Orleans because the program was not working well. There were inconsistencies in grading, and students who managed to pass the course later foundered in the credit courses. Obviously, the teachers needed to cooperate to raise the goals of the course, to establish common grading standards, and to test each student's ability to write Edited American English, concentrating on student mastery of the basic sentence. The proficiency test that was constructed is a hybrid of an empirically based error count and a broad analytic scale. The error count is based on the number and type of errors that teachers will tolerate, and the analytic scale is a measure of coherency and organization. Perhaps as important as the actual test construction is the fact that the staff reached a consensus on standards, then worked to help students meet those standards. Although the standards were higher than before, more students were now passing them. The system of checks and balances to insure that the test is a true measure of student ability has had the added effect of getting both students and teachers involved in a common goal--to master the test. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED172214

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Elizabeth F. Penfield

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Elizabeth F. Penfield
University of New Orleans

Proficiency Testing within the Basic Writing Course: Developing a Faculty Consensus on Standards

Multiple-guess, true-false, one paragraph or five, the testing of students' writing ability has always been problematic. The expedient, fast, and cheap way to do it is to use an objective test, at best a vicarious way to measure writing, one that is similar to assessing driving ability by testing knowledge of safety measures, road signs, and traffic laws. While such a test may be reliable, in that it effectively measures what it measures, it is not necessarily valid; that is it does not measure what it purports to measure--writing. Objective tests of writing in reality test editing, spelling, mechanics, vocabulary; they do not test writing. To test writing, one must ask students to write and then assess what is written. And there's the rub. The rub is particularly painful for any English department that tries to develop a proficiency test for its basic writing course and is faced with the current literature on testing. For those brave and foolhardy enough to attempt to work out a test of writing however, the rewards can be a consensus on standards that leads to a more effective course.

Recent advances in the testing of writing based on writing samples may work well in theory, but when applied to the classroom they may present more problems than solutions.[1] Diederich's holistic method is always a relative measure; NEAP's primary trait scoring requires extensive pre-testing and data analysis; analytic scales are too time consuming for a rapid reading of papers; straight error counts can be overly mechanistic. Some of these methods work best to measure the improvement of a group, others to measure individual changes, but

CS 204 866

none is readily adaptable as an exit test for a basic freshman English course, particularly when no support is handy--no budget for testing, no hot shot testing and measurement people, no extra compensation for readers. To find a test that would and could work at the University of New Orleans, we had to first figure out where we wanted to go, that is, what we wanted from the test. With that in mind, we evolved a hybrid to use as an exit examination for our basic English course. En route, we reached a consensus on standards.

To explain why we needed a proficiency test and how we devised one, I must first describe the University of New Orleans and the course. Our open admission policy each year gives us 14,000 students who come from the immediate urban area. Most of them are first generation college students, most work part or full time, and many are economically and educationally disadvantaged. Any who make an ACT score in English of 15 or below and who have written a fifty minute essay that confirms that score are placed in basic English. These students have so many strikes against them that it is no wonder so many strike out. Many read at the ninth grade level; almost all are enrolled in remedial math. While some 60% of these students are black, all of them speak a dialect at odds with Edited American English. In fact, the very act of writing is foreign to them, and they approach it with the fear and loathing many of us reserve for the IRS. Nonetheless, each fall, approximately 1000 students are in basic English, a three credit, three contact-hour course that does <u>not</u> count toward graduation. Even the grade system is different. Several years ago, we adopted Pass/Unsatisfactory grades so that students would not be penalized academically if in one short semester they could not make up for eighteen years' worth of noneducation. Ironically, however, to the basic student the grade system is one more stigma and one more indication that

the course isn't worth much.

As though all of these handicaps were not enough, we had one more: sheer size. In the fall of 1977, for instance, 1,159 students filled forty-five sections, taught by twenty-one different instructors who ranged in teaching experience from two to twenty years, in enthusiasm from lukewarm to messianic, and in attitude from dispassionate to empathetic. As a result, standards and grades fluctuated wildly. One student would receive a U from Professor X while the student's clone would receive a P from Professor Y. No matter which section, however, the goals of the course were the same, such as they were: to write at least one paragraph that was grammatically clean and mechanically acceptable within the two hours alotted for the final examination. While I suspect that paragraph was subject to varying interpretations, to say nothing of grammatically clean and mechanically acceptable, something was terribly wrong with these goals. They didn't work. As minimal as they were, in the fall of 1977, only 46% of the students met them, and of those who did and continued in college, 45% received a D or F in the next English composition course. Our basic course sorely needed revision in order to raise standards and make them more uniform. The proficiency test was an idea whose time had come.

Proficiency tests can be dangerous things--one need only think of accountability--so I wanted to make sure that whatever we developed served the purposes we intended: to raise the goals of the course, to insure consistency in grading, and to test a student's ability to write Edited American English. How we created the test was as important as what we came up with.

A few years ago, Arthur M. Cohen wrote an article called "Assessing College Students' Ability to Write Compositions."[2] On the way to designing a scale that

could be used to measure a change in the ability to write essays, Cohen found "it is possible to involve English instructors in the actual conduct of a learning study and still obtain results the researcher would find respectable."[3] While the statement may be the only documented evidence that English teachers can agree, and Cohen was writing about English instructors working together to design an experiment, the situation was certainly analogous to our trying to work together to develop a valid and reliable proficiency test. With that idea in mind, I formed a committee of eight instructors who had all agreed to tackle the task. From the beginning we all understood that what ever we came up with, we would try, and if, with a little tinkering, it worked, we would present it as a recommendation to the faculty who were teaching the basic course; if they approved, to the Freshman Advisory Committee. The last stop would be the full English faculty. We wanted everyone to understand exactly what we were proposing.

The first step was to figure out why basic students did poorly in the next English course. All of us on the committee had taught that course and had suffered the same anguish when we watched students, in spite of our efforts and theirs, students who had legitimately passed the basic course, falter, panic, and go down with a D or F. They all became croppers for the same reason: they could not write a sustained piece of prose without committing serious basic errors. No matter how great the effort, how real the voice, how touching the text, these students could not come to grips with the English sentence, a fact that earned them D's and F's. With a D or F in hand, the student was trapped, for neither repeating the course, nor advancing shakily to the next in the sequence would solve the basic sentence and verb problems. The standard composition courses were not designed for that. Our test of writing, then, had to demand some length and

had to quantify errors. A student who passed the basic course by conquering error, by internalizing the sentence sense of the language, should be in good shape to move on to the more sophisticated rhetorical problems of organization and style, concerns of the remaining composition sequence.

We set to work to narrow down error, quality and quantity, but when we started to discuss it, almost immediately we began to bicker over whether a particular example was one thing or another or even an error at all. Was "I have two brother" a problem in case or in agreement? The more important question was, did it matter? Perhaps our living in predominantly Catholic New Orleans had unconsciously affected all of us, for we finally hit upon the solution of cardinal errors and venial errors. In non-religious terms, these were errors that offended and those that irked, or, as we finally agreed, major and minor errors. Major errors were simple to define: sentence errors, such as fragments, comma splices, fused sentences; and verb errors, such as subject-verb agreement and defective verb forms. While we might squabble over nomenclature (is it a run-on or a comma splice?), we all could and would recognize a non-sentence and a screwed up verb.

Minor errors presented a greater problem because of the range of posibilities. We debated over the relative weight of spelling, punctuation, wrong words, tense shifts, agreement other than subject-verb, even linguistic niceties such as squinting modifiers and the correct use of the ablative absolute. Eventually we settled for a minimal list: ending errors (-s, 's, adjectivel -ed omitted), wrong words, omitted words, illiterate punctuation, and illiterate spelling [4] These last two categories call for further explanation. We agreed that a student must have a working concept of the purpose of punctuation; beyond that, a handbook could straighten out problems. A comma plunked between a subject and verb, a

semicolon poised after an introductory phrase, a hyphen fracturing a one syllable word doesn't give much for a handbook to build on. They are illiterate uses of mechanics. The same principle applies to spelling. If college comes out colidge or calm turns into kam, the student cannot use a dictionary and therefore has no resource to draw upon. Illiterate spelling, then, is misspelling that cannot be cured by the dictionary.

Having determined what, we turned to how and how many. Since the next course called for a fifty minute in-class paper of 350-400 words, we decided that 300 words would be the goal in the basic class. Given the fact that impromptu themes terrify everyone and that basic writing students have trouble organizing their experiences into a paper,[5] we settled on announcing the topics before the exam and then each instructor selecting one of them on the test day. That way students would have a chance to think about the topics and organize their ideas, but at the same time would not be reeling off memorized themes. Armed with the goal of 300 words and the constraint of fifty minutes, we tried to figure out our tolerance for error, finally agreeing on no more than a total of two major errors (no more than one sentence error of the same kind); and no more than a total of eight minor errors (each error a separate error except for misspellings of the same word).

In a test that is scored by an error count, the number and the weight of the error are important, not the type. Thus we avoided squabbles over whether behin, as in "I work behin the counter," was an illiterate spelling or non-word; either way, it was minor error--one out of an allowable eight. To call behin an error is indisputable, an objective assessment, but in order to make the test as unbiased as possible we took further precautions: neither the student nor the student's teacher's name appeared on the paper and each paper had two

independent readers in order to check on consistency in scoring, a third reader in case of disagreement.

Aside from being fairly free from errors, a passing paper had to fulfill three other requirements: it had to be of sufficient length, on the assigned topic, and coherent. A long paper received some leeway; a 250 word paper passed if it was clean. But no matter how clean the paper, it did not pass if it lacked a major idea that was supported by details arranged in some logical order. This requirement toughened the test and made it less mechanical, but we worried about its being too tough. Our fears of organization and development doing in students vanished, however, when we gave the test a dry run, dry in that it had no effect on the student's final grade. We found that what we had suspected was true: students whom we evaluated subjectively as passing students passed the test, and, more to the point, almost all the students who could control error, who could proofread, wrote organized and developed papers.

We had developed a proficiency examination that we could use as an exit test for the basic course, but we had to build an escape route. Anyone can have a bad day, and good developmental writers are particularly susceptible to them, a fact Mina Shaughnessy notes in her Errors and Expectations. With that in mind, we designed a system for challenging the results of the test. When the grade sheets were returned to the instructors, each would go over the results carefully. If a student who was clearly passing the course failed the exam, the instructor would turn in the student's complete folder and an explanatory note. The folders would be reviewed independently by two readers, a third in case of a dispute. If two readers said pass, pass is what the student received in the course.

The reverse was also true. If a student who had missed every other class and had written failing papers miraculously produced a passing examination, the

instructor would challenge the proficiency grade, turning in the student's complete folder and a brief note explaining the problem. If the two readers said unsatisfactory, unsatisfactory is what the student received in the course.

After testing the test, modifying it slightly, getting the support of the teachers of the course and the imprimatur of the department, we adopted it. And we are happy. The standards of the course are higher, uniform, objective. Students know from the start of the semester exactly what they must do in order to pass, and both students and teachers work toward that end. The fact that we have reached a consensus on standards and that that consensus has been defined, quantified, and made objective has made the course real to the students who are in it. The course still does not count toward graduation, the students in it know they are in a basic course, and the grade system is still different (U/P), yet the stigma associated with these differences has been diminished to the point where it is no longer significant. Too often this stigma "reinforces the attitudes [students] have developed during the last twelve years: they expect to learn nothing and assume that, since they are inferior, they will fail."[6] Clarifying standards and applying them uniformly eliminates this problem by making the course's goal concrete and progress toward it visible. Students want to succeed and do. Before we adopted the competency test, 46% of the students passed the course; with the exam, even though standards were raised significantly, 48% passed.

Perhaps the most beneficial impact of the proficiency test is that it has given everyone--student, teacher, administrator--an enemy: the test. The enemy exists, is tangible, and can be conquered. Teachers and students are on the same side, working to win, to pass the course. The students no longer can accuse the faculty of capricious grading, and any lingering doubt in either the teacher's or students' minds can be answered by a challenge. For the students, the test has

made the course real, giving them concrete measurements of progress as well as of success or failure; for the faculty, the test has raised morale. A semi-dormant committee composed of all the basic staff has come alive. Because the test is a test of writing, faculty discussions focus on the process of writing and the patterns of error. Trivial surface features are treated as such and the larger problems of coherence are tackled directly. In the classroom, problems of grammar and usage are treated within the context of the students' writing; out of the classroom, the staff meets frequently to exchange ideas, trade techniques, and help each other help students.

The test we ended up with is a hybrid combination of an empirically based error count and a broad analytic scale. The error count is based upon the number and type of errors teachers can tolerate, and the analytic scale is a gross measure of coherency and organization. But perhaps as important as developing the test itself is the fact that the staff reached a consensus on standards and then worked with the students to help them meet those standards. As we developed a test for the students, we tested ourselves, constantly rethinking and redefining our own objectives and values. I think we passed.

NOTES

[1] Various methods of measuring and evaluating writing are thoroughly described in Charles Cooper and Lee Odell's Evaluating Writing: Describing, Measuring, Judging (Urbana: NCTE, 1977).

[2] Research in the Teaching of English, 7 (Winter, 1973), 356-372.

[3] Cohen, 369.

[4] Roberta Mathews' "The Evolution of One College's Attempt to Evaluate Student Writing," Basic Writing, 4 (Spring/Summer 1978), 63-70, also describes an error count, but the list is less well defined than ours and, when the essay is scored, errors are not tallied separately but marked on the paper, a method that introduces possible bias.

[5] Sara Sanders and John H. Littlefield are able to demonstrate the dangers of the impromptu theme and the advantages of announced topics in "Perhaps Test Essays Can Reflect Significant Improvement in Freshman Composition." Research in the Teaching of English, 9 (Fall, 1975), 145-154.

[6] Vernon E. Lattin, "A Program for Basic Writing," College English, 40 (November, 1978), 313. Lattin argues in favor of a competency-based, regimented credit course. In doing so, he discounts the non-credit course without considering the benefits it can gain from competency testing.